WORKING MEMORY CAPACITY INFLUENCES COGNITIVE FLEXIBILITY: UNDERSTANDING THE CONTROL DILEMMA THROUGH STROOP

by

Katherine F. MacNamee

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Psychology

The University of Utah

May 2016

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Katherine F. MacNamee**

has been approved by the following supervisory committee members:

Jason M. Watson	, Co-Chair	**3/4/2015** Date Approved
David Strayer	, Co-Chair	**3/4/2015** Date Approved
Jeanine K. Stefanucci	, Member	**3/4/2015** Date Approved

and by **Lisa Aspinwall**, Chair of

the Department of **Psychology**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

We sought to determine whether Working Memory Capacity (WMC) predicts when an individual will exert or withhold cognitive control when faced with a control dilemma. We employed a high-congruency variation on the Stroop task to maximize conflict between automatic and controlled processing, and manipulated task instructions between participants to emphasize the importance of exerting cognitive control or convey typical speed/accuracy instructions. A 2 (trial type) x 2 (instructions) x 4 (WMC [quartiles]) analysis of variance (ANOVA) revealed an interaction, in which instruction manipulations failed to affect the proportion of errors made by low- or mid-span individuals. High spans, however, made a lower proportion of errors when warned of task pitfalls than when not. Regression analyses suggested that, when warned of the pitfalls of relying on automatic processing, WMC and proportion of Stroop errors exhibit a negative, linear relationship. However, when the nature of the need to exert control was not explicit, a curvilinear pattern was observed. Those with high WMC appeared to strategically withhold control, relying instead on automatic processing. This led them to make a higher proportion of Stroop errors. Response latency data suggested that lower-mid-spans were most rigid in their exertion of control, while high spans were especially flexible across instruction conditions. These data suggest a higher WMC allows for increased cognitive efficiency of cognitive control exertion across varied contexts. These

results could be the product of an increase in cognitive resources allowing for better metacognitive abilities in those with high WMC.

To Susan and Jay,

for teaching me to value education in all its unexpected forms.

CONTENTS

ACKNOWLEDGEMENTS

Special thanks to Jason Watson for his invaluable and ever-optimistic support throughout this process; your mentorship has forever changed the scientist I am and will become. Thanks also to Dr. Stefanucci and Dr. Strayer for their guidance these last two years. Thank you to my research assistants, particularly David Alonso and Maryanne Emery, for their incredible dedication to quality data collection; Stroop-ing for hours can be a tiresome task, but you remained gracious and hardworking all the way through. Finally, thank you to Tyler, whose words of encouragement and belief in my work made even the difficult moments manageable.

INTRODUCTION

Cognitive control is defined as the regulation of cognitive processes, such as reason and task-flexibility. It is influenced by many factors, both personal and environmental (McVay & Kane, 2009; Prakash, Hussain, & Schirda, 2015; van Steenbergen, 2015), but working memory capacity (WMC), in particular, has emerged as a strong predictor of one's ability to exert cognitive control. Furthermore, research suggests that individuals with high WMC (high spans) are better able to exert cognitive control than those with low WMC (low spans; Kane, Conway, Hambrick, & Engle, 2007; McVey & Kane, 2009).

Interestingly, prior studies have also shown that situations exist in which high spans perform similarly to low spans (Miller, Watson, & Strayer, 2012; Watson, Miller, Lambert, & Strayer, 2008). This pattern seems to occur most noticeably during control dilemma tasks, which pit automatic and controlled processing against one another. Such tasks allow individuals to choose whether they exert control — and thus achieve better task performance — or strategically withhold it, thereby conserving cognitive resources for later use (Kahneman, 2011). Specifically, high spans fail to differ from low spans during high-conflict control dilemma tasks in which the importance of suppressing an automatic response during conflict is not made clear (Watson, Miller, Moffitt, Strayer, & Lambert, 2013). Still, psychologists have predominantly studied this phenomenon using motoric and sub-cortical automaticity, so high and low spans may only behave similarly during tasks that tap into highly potent automatic processes (e.g., antisaccade or Simon). Moreover, it remains unclear how those with intermediate WMC (mid-spans) behave in the face of a high-conflict control dilemma, and whether WMC is, indeed, the driving force behind the differences in control exertion observed at the extreme ends of the WMC spectrum. Thus, the present study sought to answer the following questions:

1) Are previously observed control exertion patterns for differing WMCs (e.g., Watson et al., 2013) domain free? That is, can such patterns be found when the automatic responses in question are *learned* and therefore deal with higher order functions (e.g., word reading; Stroop), rather than when the automaticity is subcortical or motoric (e.g., button presses; Simon)?
2) In addition to finding differences between the extreme ends of the WMC spectrum, as has been shown in previous studies (e.g., Hiebel & Zimmer, 2015; Storbeck, Davidson, Dahl, Blass, & Yung, 2015), will we find curvilinear (rather than linear and dichotomous) exertion patterns when accounting for the entire range of WMC levels? In other words, how can the behavior of the middle 50% of the WMC spectrum support or alter our theories as to the reason high spans sometimes differ in behavior from low spans?

Answering these questions will allow us to better understand how the availability of cognitive resources (as measured by WMC) interacts with context to affect an individual's decision to exert or withhold control. This should allow us further theoretical insight into the relationship between WMC and controlled attention.

Historically, working memory models involved a storage component, as well as a controlled processing component (see Miyaki & Shah, 1999, for review). In the last decade, however, the field of cognitive neuroscience has widely begun to recognize that WMC can be reconceptualized as the ability to allocate resources to executive attention and control (Kane et al., 2007). In light of this theoretical shift, many recent theories of attention assert that high WMC is associated with increased capacity for exerting cognitive control when achieving a goal or manage sources of conflict or interference (Shipstead, Lindsey, Marshall, & Engle, 2014). A good deal of work also suggests that WMC predicts performance in processes, such as stereotype regulation, action monitoring, multitasking, and inattentional blindness, (Lambert, Seegmiller, Stefanucci, & Watson, 2013; Miller et al., 2012; Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013; Seegmiller, Watson, & Strayer, 2011, respectively). Thus, it stands to reason that tasks which tap into higher level processes by pitting controlled and automatic processes against one another also give advantage to those with high WMC (Engle, Kane, & Tuholski, 1999). In fact, psychologists use many such control dilemma tasks when studying the exertion of control (e.g., Simon, Eriksen flanker, antisaccade, Stroop, and value-directed memory tasks). Observing how individuals with different WMCs respond to these tasks may help us to determine the mechanisms that allow for flexible transitions between automatic and controlled processing.

However, control dilemma paradigms traditionally create situations in which the exertion of control leads to notably increased success, thereby placing an implicit premium on controlled processing. Relying only on controlled processing to accomplish one’s goals makes little sense when we consider that strategically withholding control in tasks of low importance allows us to conserve cognitive resources and benefit later on (Kennet & Wurm, 2014; Meiring, Subramoney, Thomas, Decety, & Fourie, 2014; Ward

& Mann, 2000). It is not necessarily obvious in day-to-day interactions who among us has a high or a low WMC. Therefore, modifying established paradigms to increase the likelihood of relying on automatic processing serves simultaneously to bring participants down from ceiling (Kane & Engle, 2003) and to create a more realistic scenario in which to investigate control exertion. Moreover, paradigms that are somewhat ambiguous in their required level of control exertion (i.e., high-conflict control dilemma tasks) may make it easier to manipulate whether an individual chooses to exert control.

The possibility of manipulating exertion is supported by Goschke (2003), who asserted that controlled processing can be modulated by conscious intent, as well as activated automatically. Data from Stawarczyk, Majerus, Catale, and D'Argembeau, (2014) suggest that the exertion of cognitive control can be both facilitated and hindered by bottom-up, or stimulus-driven processes. These findings imply that explicit instructions aimed at altering the perceived importance of controlled processing might prompt participants to exert more control (see Watson et al., 2008), but the effects of instructional changes might differ for different WMCs. Conway and Engle (1994) theorized that differences in WMC actually reflect individuals' varying abilities to inhibit irrelevant stimuli and attend to task-relevant information, rather than reflecting differences in their absolute capacity for holding information. Thus, it seems likely that differences in WMC – in addition to explaining differences in exertion ability – also change *why* and *when* individuals choose to exert control when presented with processing conflicts (Engle & Kane, 2004; Watson et al., 2013).

While the current literature addresses some of the ways in which the exertion of cognitive control can be influenced, most studies of control exertion are designed using paradigms that require reliance upon strongly automatic processes or strongly controlled attention to succeed. However, a growing number of theories surrounding cognitive control suggest that it is no longer advantageous to create a binary relationship between automatic and controlled processing. This is where control dilemma tasks, such as Stroop, become important.

The Stroop task requires participants to suppress the automatic response of reading a color word (e.g., RED), and instead use controlled attention to verbalize the color in which the word is written (e.g., blue). In congruent trials, the color word and the font color are the same (e.g., **RED**), so relying on the automatic response of word reading does not decrease an individual's level of performance. By contrast, incongruent trials contain stimuli written in a different color than the word itself (e.g., **RED**), so reliance on automaticity will result in error. We believe that further knowledge of cognitive control might be best acquired by using variations of Stroop that maximize conflict by containing a high ratio of congruent (automaticity-friendly) to incongruent (control-necessary) trials. This would increase conflict and task ambiguity, because when the already-easier option of falling back on automatic processing leads to a fairly acceptable success rate, the habit of relying on automaticity becomes more likely to occur. When the perceived benefit of using either automatic or controlled processing becomes virtually equivalent, participants' reliance on context when deciding whether to exert or withhold control should increase. In the present study, examining differences in both WMC and context as they relate to control exertion should provide a better understanding of the relationship between WMC and cognitive control.

Furthermore, in varying the prepotency of the task from subcortical activation (such as that needed for anti-saccade), and even going so far as to remove the manual motoric aspects of the task (as needed for Simon), we can use the Stroop task to test the circumstantial limits of the control dilemma. That is, we can learn whether the relatively lower prepotency of word reading is still strong enough to cause those with available cognitive resources to rely on automaticity, or whether the cost of exertion is close enough to the cost of relying on automaticity that those with available cognitive resources continue to exert control in all contexts. Both the decreased prepotency of the task and the high ratio of congruent trials serve to limit the influence of task-inherent cues that might facilitate control exertion. This will allow us to more easily manipulate explicit contextual cues that encourage exertion — at least in those with cognitive resources to exert (i.e., high spans).

To create explicit contextual changes and determine their influence on control exertion, we altered our task instructions to suggest either high or ambiguous importance of exerting control. Similar techniques were used by Watson et al. (2008, 2013) to elicit control exertion in the Simon and Saccade tasks. It may be that those with high WMC are willing to use their resources to overcome the automatic response of word reading when informed that high accuracy for incongruent (control-necessary) trials is the goal of the task. The question then becomes whether the abundance of resources in high WMC individuals would be worth spending on suppressing the automatic word reading response, especially in a high congruency situation.

If high spans exert control when warned of the importance of incongruent trials, but rely on automaticity when unaware, we will see results similar to what prior studies (Watson et al., 2013; Watson, Bunting, Poole, & Conway, 2005) have found. If, however, contextual changes have no effect, we may see high spans exert control regardless of context and suppress the automatic response. This might occur because they have resources to spare, or because word reading is not a potent enough automatic response. However, Watson et al. (2005) asked participants to complete a Deese-Roediger-Mcdermott (DRM) paradigm, which like Stroop contains verbal stimuli. For this study, the authors also manipulated instructional warnings — in this case to explain (or not) the risk of false memories — and used participants' WMC as a predictor. They found that instances of false memory were reduced only for high spans in the warning condition, but that no differences presented between high and low spans in the no-warning condition. Although automatic spreading activation in the DRM and the automatic response of word reading in Stroop differ in prepotency, both involve verbal stimuli. Based on results from the Watson et al. study, we predict that high spans will show greater advantage as a result of explicit warnings, compared to low spans. Conversely, we predict that the no-warning condition will indicate no significant differences between the upper and lower extremes of the WMC spectrum. However, even if the pattern of high span exertion remains similar to what has been observed, it is still unclear how those with intermediate WMC might respond to varying contexts.

Oddly, few control exertion studies, if any, examine working memory on a continuous scale, choosing instead to focus on the extreme ends of the WMC spectrum. This methodological choice does not account for a large portion of the population, and could mask important information, which would further our understanding of how individual differences in WMC affect cognitive control. If, for instance, there is a linear relationship between WMC and Stroop performance, we should find that mid-spans perform better relative to low spans, and worse relative to high spans. On the other hand, high spans may choose not to exert control unless warned, because their ample cognitive resources are not fully drained by the task, and they therefore have resources left over, which enables them to learn — fairly early on — that the congruency ratio allows them to rely on word reading and still succeed. Simply put, high spans might just be better than others at maximizing cognitive efficiency. Should this be the case, the mid-spans would theoretically exert control in both contexts, yielding a curvilinear pattern in Stroop error rate. This is because although in relation to low spans, mid-spans have ample control to exert, they may not have enough resources to successfully complete a task while simultaneously capitalizing on their metacognitive abilities. Without access to sufficient metacognition, they could perform well, but would do so at the cost of efficiency. Thus, with mid-span response patterns adding potential nuance to the data, we can better infer reasons as to the exertion (or lack thereof) on the part of high WMC individuals, based on the similar or differing behavior of the intermediate group.

METHOD

Participants

Participants consisted of male and female undergraduate students (N = 180, ages 18-40) recruited from a mountain west university. Each student received course credit for their participation. All possessed normal or corrected-to-normal vision, as well as normal color vision. Participants were tested in individual rooms for a single session, which lasted no more than an hour. We recruited the first 14 participants based on their working memory capacity (WMC), as determined by operation span task scores from a separate experiment. These participants were chosen, because they had scored in either the lower or upper quartile, according to thresholds determined by Unsworth et al. (2005). The remaining 166 participants were recruited from the psychology department participant pool. Of these, 4 participants were excluded from analyses due to computer errors, 4 due to failure to maintain accuracy for reading span sentence verification, and 3 due to average response latencies greater than 1250 ms in Stroop. Thus, we analyzed data from a total of 169 participants.

Materials

E-Prime. To present both the reading span and the Stroop task, we used the 2003 version of E-prime software (PST, Inc.).

Reading span. We asked participants to complete an automated reading span task (as described in Unsworth et al., 2005). The program displayed all stimuli in black, Times New Roman, 24pt font on a white background. The testing portion consisted of a maximum 15 trials (three of each set size), with set sizes ranging from 3 to 7 letters. A sentence verification (length of 5-10 words) required participants to make sense/nonsense judgements between the appearances of each letter.

Stroop. Our version of this task included only the color words RED, GREEN, and BLUE, with the same three corresponding font colors. The stimuli contained a proportion of 75% congruent trials (e.g., RED written in red font) and 25% incongruent trials (e.g., RED written in green font) with the intent of maximizing participants' reliance on the automatic process of word reading (see Kane & Engle, 2003 for a review). Within the congruent trials, we used an equal number (25% each) of the words RED, BLUE and GREEN. Of the incongruent trials, we used an equal number of each possible variation (RED, RED, GREEN, GREEN, BLUE and BLUE). All trials were presented in Times New Roman 24pt font in one of the three font colors and always on a black background. The testing portion consisted of three blocks, each containing 168 randomly presented trials.

Procedure

After obtaining informed consent, we began the experiment, which was separated into two tasks: the automated reading span and the Stroop.

Reading span. Participants completed a computerized reading span task as a measure of WMC. The task asked participants to simultaneously memorize a series of letters, which appeared on the screen in between sentences. Participants determined whether each sentence was either sensical (e.g., The dog likes to go for walks with his owner.) or nonsensical (e.g., The colander likes to go for walks with his owner.), and responded with “True” or “False,” respectively. We gave participants written instructions within the E-prime platform, as well as a verbal reminder to read the instructions carefully before starting. We then left participants alone in the testing room as they complete the reading span, which consisted of practice for letter recall, practice for sentence verification, and practice performing both together before the experimental trials began. Upon completion of the testing portion, participants notified the researcher, who recorded the five reading span output scores and set up the Stroop task, while the participant rested (approximately 2-5 min).

Stroop. Participants sat directly in front of the monitor and close to the tabletop microphone, which recorded vocal response latencies (RTs) using an SR box. The researcher sat beside the participant and used the keyboard to code participant response accuracies and SR box errors. The possible codes were as follows: c- correct, clean RT; i – incorrect, clean RT; s – correct, SR box error; o – incorrect, SR box error; u – unknown/no response due to SR box error. The researcher also read aloud all instructions that appeared on the screen. Before beginning each of the testing blocks, participants completed a set of 12 practice trials.

In the “warning” condition, we gave participants instructions that warned of the high congruency proportion, and explained that reliance on reading the color word during congruent trials will lead to poor performance when confronted with incongruent trials. Instructions read:

> In the following trials, you will be shown color words (e.g., BLUE) written in various font colors (e.g., red). Please say aloud the color in which the word is written, rather than the word itself. Remember to be as fast and as accurate as possible. You may find in many trials, the color word (e.g., RED) will match the font color (e.g., red), making it easy to respond with the correct font color. Remember, these are filler trials meant to distract you

> from the given task of font color naming, and they allow you to rely on word reading to give the correct answer. This may cause you to perform poorly when the color word and font color don't match (i.e., BLUE is written in red or green font), even though these nonmatching trials are the critical trials in which we are interested.

In the “no-warning” condition, we gave only the speed/accuracy instructions, which consist of the first three sentences of the warning condition text. The program reiterated instructions at the start of each of the three blocks. Upon completion of the testing blocks, participants were debriefed, given credit, and dismissed from the experiment.

RESULTS

We excluded all trials with a length greater than 1250 ms or less than 250 ms from analyses. This removed fewer than 2% of trials from subsequent analyses, and did not include trials in which the SR box was coded to have erred. We excluded trials containing SR box errors from response latency analyses, but not from error analyses.

Stroop Error Rate

ANOVA. We first examined only the upper and lower quartiles of WMC in a 2 (warning condition) x 2 (trial type) x 2 (WMC span) mixed model analysis of variance (ANOVA) to determine whether the error proportions observed in high and low spans in the present study resemble those found in previous work (e.g., Kane & Engle, 2003). Analyses revealed a three-way interaction, $F(1, 81) = 6.60$, $p = .012$, $\eta_p^2 = .075$. Additional analyses indicated no significant differences between high and low spans for incongruent trials in either warning group, *p*s $\geq .21$. In contrast, within incongruent trials, high and low spans differed significantly in the warning condition $p > .001$, but failed to differ in error rate when no warning was given, $p = .24$.

To next determine whether all four WMC quartiles differed from one another across various groups, we submitted the data to a 2 (warning condition) x 2 (trial type) x 4 (WMC span) mixed model ANOVA. See Table 1 for means. The data revealed a significant three-way interaction, $F(3, 161) = 5.68$, $p = .001$, $\eta_p^2 = .096$. First, within congruent trials, we found no differences in Stroop error rate across either WMC span or warning condition, Fs ≤ 1.12, ps $\geq .40$. In contrast, significant differences for incongruent trials for both WMC and warning condition were apparent, $F(3, 165) = 19.73$, $p < .001$, $F(1, 165) = 4.84$, $p = .029$, respectively.

Within incongruent trials in the warning condition, only high spans and upper-mid-spans failed to differ significantly from one another in proportion of Stroop errors made, $p = .63$, all other ps $\leq .001$. Low spans made the highest number of errors, followed by lower-mid-spans, then upper-mid-spans and high spans (Figure 1). See Table 1 for means.

For incongruent trials in the no-warning condition, a different pattern emerged. Low spans made significantly more errors than lower- and upper-mid-spans ($p = .007$, $p = .001$, respectively), but failed to differ in any meaningful way from high spans ($p = .24$). Upper and lower-mid-span error rates were not significantly different from one another ($p > .99$), and high spans did not differ significantly from any of the other groups, all ps $\geq .24$ (Figure 1). We think it important to note that only high spans' error rates were significantly affected by warning condition, $p = .001$, all other ps $\geq .34$.

Generally, we believe these accuracy data indicate an increased flexibility in high spans' control exertion, compared to those with low or intermediate WMC. To further investigate the differing patterns of error rate across the full WMC spectrum, rather than arbitrary bins, we turn to regression analyses.

Regression. For the warning condition, we used a linear regression to determine whether working memory capacity (WMC) predicted Stroop performance when instructions explained the importance of incongruent trials to participants. We found a significant, negative relationship between WMC and proportion of errors made, ß = -.77, $t(84) = 21.46$, $p < .001$, with an increase in WMC predicting a decreased proportion of Stroop errors (Figure 2). WMC also explained a significant proportion of variance within Stroop error rate, $R^2 = .59$, $F(1, 84) = 123.07$, $p < .001$. This suggests that, when told control exertion is necessary for success, those with higher WMC were both willing and able to allocate cognitive resources to the task. Those with lower WMC, on the other hand, seemed less able to exert cognitive control.

For the no-warning condition, a linear regression tested whether the pattern observed above could be seen in the absence of task-related information aside from typical Stroop instructions. The model fit was marginally significant, ß = -.21, $t(83) = 3.9$, $p = .052$, suggesting a somewhat negative relationship between WMC and proportion of Stroop errors made. We then applied a quadratic curve estimate to the data, and found that the quadratic model accounted for a significant proportion of variance within Stroop errors, $R^2 = .12$, $F(2, 82) = 5.20$, $p = .008$ (Figure 3). This model predicted that those with intermediate levels of WMC made fewer errors than individuals on either end of the WMC spectrum. The relatively larger number of errors made by higher spans suggests that, when not warned, those with high WMC strategically withheld control, leading them to commit a greater proportion of errors.

Reaction Time

We excluded all incorrect trials (approximately 8.6%) from response time (RT) analyses. Again, we began with a 2 (warning condition) x 2 (trial type) x 2 (WMC span) mixed model ANOVA, with trial type as a repeated measure, to examine differences between high and low spans in response latency. The analysis yielded an interaction between trial type and WMC span, $F(1, 81) = 18.73$, $p < .001$, $\eta_p^2 = .19$, wherein no differences were observed between high and low spans for congruent trials, $p = .54$, but high spans had a significantly faster naming speed than low spans during incongruent trials, $F(1, 84) = 7.12$, $p = .009$.

ANOVA. In order address our research question of how results change when we include the middle 50% of the WMC spectrum, we next submitted the data to a 2 (warning) x 2 (trial type) x 4 (WMC Span) ANOVA. Means for all conditions and sub groups are represented in Table 1.

We found a main effect of warning condition on average response latency, $F(1, 161) = 4.3$, $p = .040$, $\eta_p^2 = .026$, wherein participants who were warned responded slower than those who were not, suggesting, generally, that the warning condition encouraged the exertion of control.

Though no main effect of WMC was present, $p = .13$, we found a significant interaction between WMC and trial type, $F(3, 161) = 7.65$, $p < .001$, $\eta_p^2 = .13$. This was due to congruent trials showing no differences between spans, $F(3, 165) = .91$, $p = .44$, whereas RTs during incongruent trials changed significantly across spans, $F(3, 165) = 3.26$, $p = .023$, $\eta_p^2 = .056$. Within incongruent trials, pairwise comparisons revealed that high spans were faster to respond than lower-mid-spans, $p = .020$ (Figure 4).

No further significant results were observed, $F\text{s} \leq 1.88$, $p \geq .13$. And the absence of a three-way interaction for participants' average RT stands in contrast with the error rate data, for which both the traditional (2x2x2) and the full quartile (2x2x4) ANOVAs yielded significant three-way interactions.

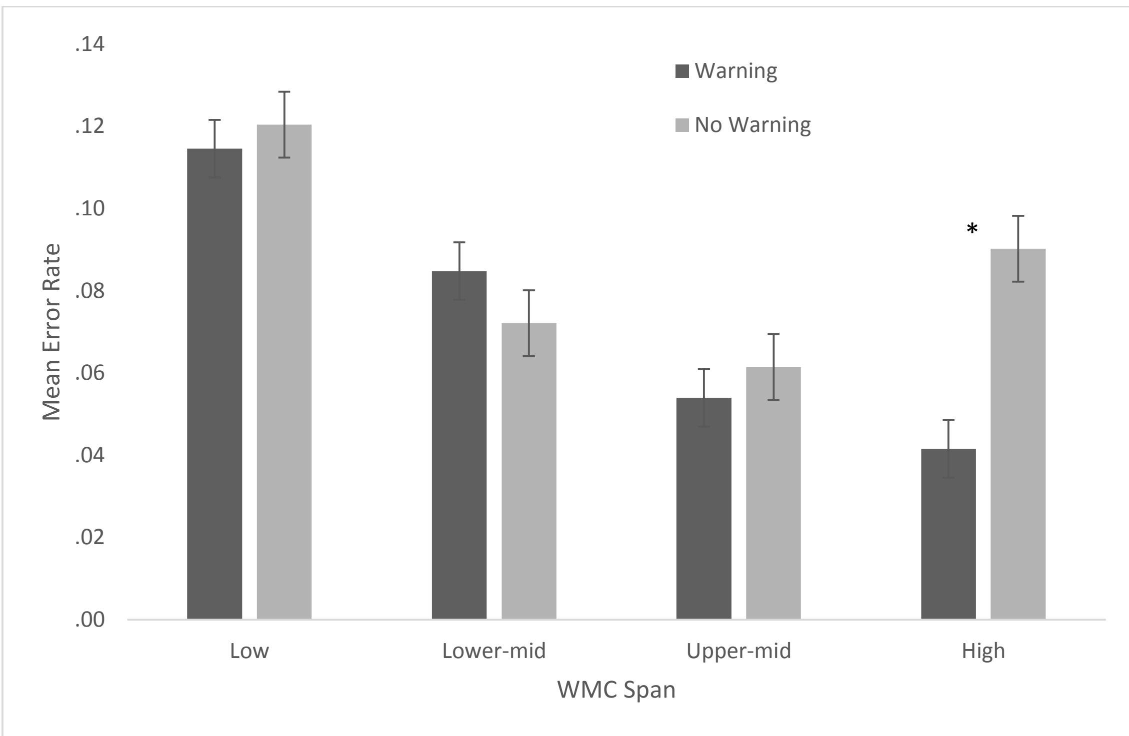

Figure 1. Average Stroop error rate for incongruent trials by WMC quartiles in the warning and no-warning conditions. Error bars denote standard error of the mean.

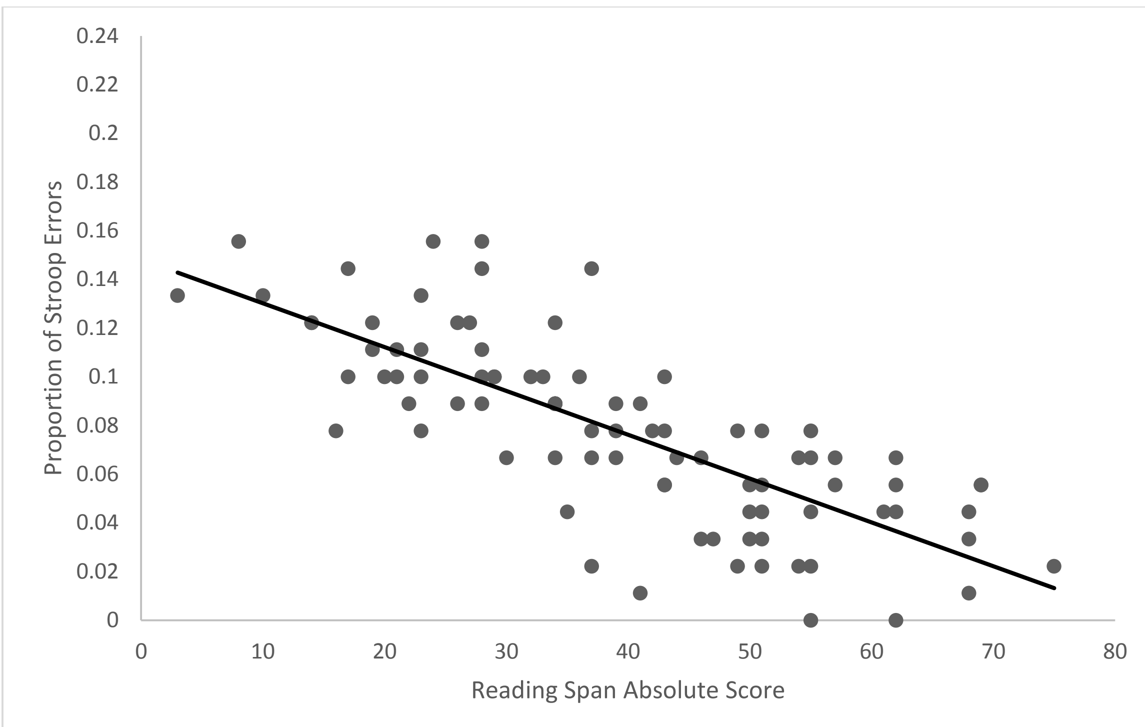

Figure 2. Stroop error rate by WMC in the warning condition.

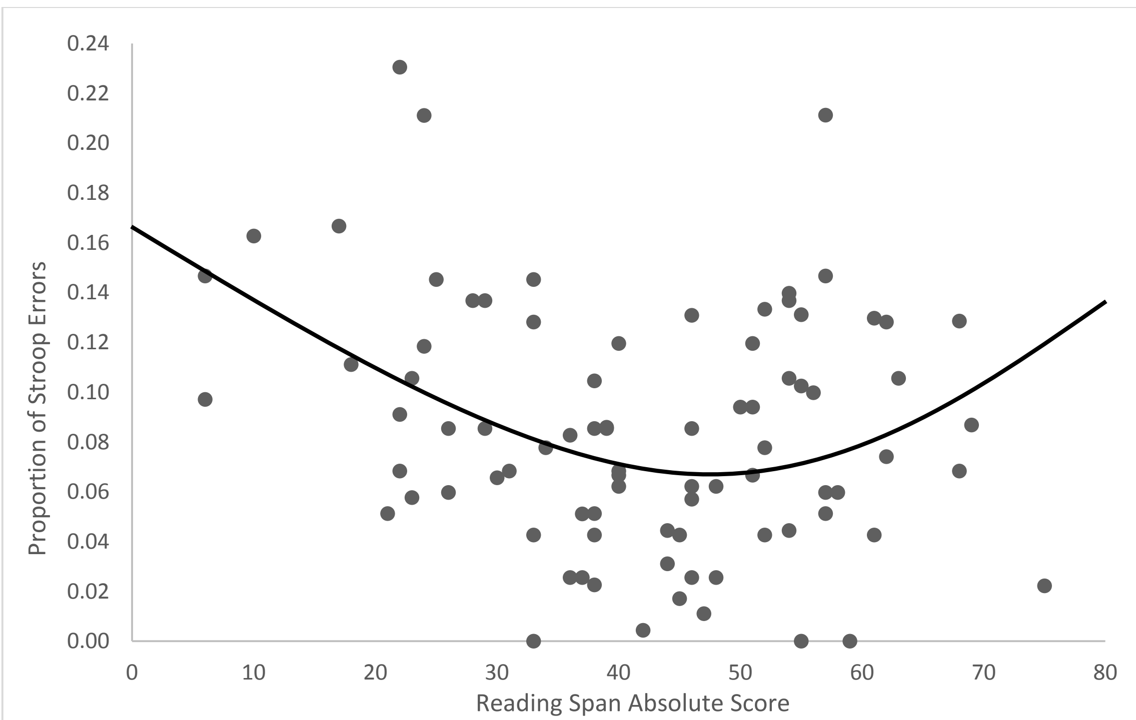

Figure 3. Stroop error rate by WMC in the no-warning condition.

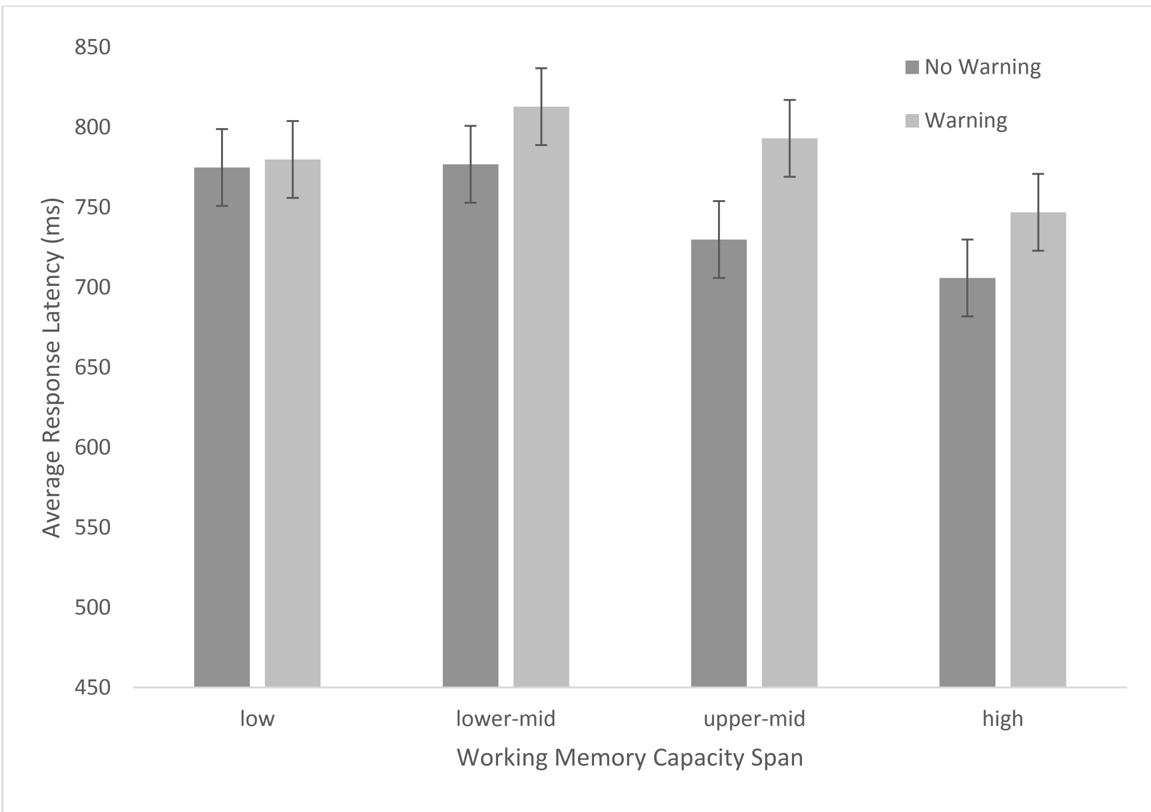

Figure 4. Average response latency within incongruent trials by WMC and warning condition. Error bars denote standard error of the mean.

Table 1

Average Proportion of Stroop Errors and Average Response Latencies (ms) for Different WMC Spans across Warning Condition and Trial Type

	Trial type							
	Congruent				Incongruent			
	Latency		Error rate		Latency		Error rate	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
No warning								
Low	596.69	68.21	.005	.0083	767.89	75.83	.12	.052
Lower-mid	619.93	67.40	.006	.012	772.13	72.46	.072	.037
Upper-mid	598.12	85.97	.0048	.0061	725.30	102.13	.061	.038
High	577.43	65.14	.0031	.0041	705.24	89.03	.09	.051
Warning								
Low	608.69	64.51	.0017	.0028	769.62	78.88	.11	.022
Lower-mid	617.20	105.86	.0027	.0036	793.45	126.16	.085	.027
Upper-mid	630.43	84.39	.0021	.0038	779.76	114.04	.054	.024
High	616.26	86.02	.0010	.0027	742.60	107.17	.042	.023

DISCUSSION

We measured WMC using an automated reading span task and asked participants to complete three blocks of a high-congruency variation of the Stroop Task. Half of the participants were given established speed/accuracy instructions for the Stroop, while the other half were told that incongruent trials were the focus of our study, and that relying on the automatic response of reading could be detrimental to their performance. The following discussion will address whether our data provided evidence for a domain-free control dilemma, what the response patterns from individuals with intermediate WMC add to theory, and what the underlying mechanisms of our results might be.

In the warning condition, the present study replicated previous work, in which high and low WMC inversely predicted Stroop performance in high-congruency task variants (e.g., Kane & Engle, 2003). Kane and Engle's study also used a 75% congruency version of Stroop and they discovered that WMC and Stroop error rate were negatively correlated, but their instructions were quite similar to those used in our warning condition. Thus, while they were very much correct in their assertion that WMC and Stroop performance were linearly related, it seems their results were most likely due to the warning embedded in the instructions they used prior to having participants begin the task.

We, too, established a clearly linear relationship between WMC and proportion of errors made on incongruent trials, and found that WMC related inversely to Stroop error rate when the context reinforced the task goal (i.e., when participants were warned). The addition of mid-spans to the analyses further supported this relationship, in that those with roughly average WMC fit soundly within the expectations of previous work, making a lower proportion of errors compared to low spans, but a greater number compared to high spans. Thus, the presence of an explicit instruction for the need to exert control makes it such that no strategic conservation of cognitive resources occurs anywhere on the WMC spectrum. Otherwise said, individuals in the warning condition performed to what was roughly the best of their ability. In the absence of a warning, however, a curvilinear pattern emerged, wherein higher spans began to make more errors than had been observed in the warning condition, and began to behave more like low spans, likely due to increased reliance on automaticity. This is where the present study deviated from previous works, and these no-warning data provide crucial insight into the relationship between WMC and control exertion flexibility.

One such insight relates to our question of whether or not the habit of word reading is practiced enough to produce a control dilemma. Furthermore, it appears that the automaticity associated with reading was, in fact, strong enough to cause sufficient conflict and prompt high spans in the no-warning condition to rely on automaticity. The differential response patterns of behavior observed across conditions in high WMC participants would suggest that the task did, in fact elicit a control dilemma. Thus, we conclude from the data that a learned automatic response, such as reading, which is consistently practiced, becomes habit enough to cause cognitive processing conflicts. This is at least true for high-congruency variations of the task.

In addition to demonstrating a control dilemma in Stroop, the curvilinear pattern of error rates during incongruent trials in the no-warning condition allows us to answer our second question: how including mid-spans in analyses advances our understanding of WMC and its relationship to control exertion. We believe our data provide considerable evidence for individual's cognitive resource availability dictating not only how well, but also *when* that individuals can/will exert control. In particular, the data allow us to focus our theories on high span behavior as anomalous. Without demonstrating that mid-spans continue to exert control, regardless of context, it would be difficult to argue that changes in high span performance are the exception. That is, prior to including intermediate WMC, one could have asserted that high spans' ample resources made them especially flexible compared to other WMCs just as easily as they might have argued that low spans' lack of resources made them especially rigid. The fact that only high spans altered their behavior when not warned to strategically withhold cognitive control suggests that cognitive efficiency is not the norm. Furthermore, only at a certain level of executive functioning can an individual allocate their resources in such a way that they can complete the task while simultaneously keeping in mind a bigger-picture perspective of their goal. Therefore, we interpret our curvilinear error data as indicating better metacognitive abilities within high spans. Generally speaking, it seems that the addition of mid-spans has allowed us to narrow down potential underlying mechanisms for behavioral patterns.

The response latency data build nicely upon the evidence provided by our error data, and ultimately indicate that high spans could possess exceptional cognitive control exertion flexibility. Longer RTs for incongruent trials compared with congruent trials serve as a reassurance that Stroop interference occurred, and that interference required increased cognitive control. Longer response latencies in the warning condition compared to the no-warning — specifically for incongruent trials — suggest that the warning did, in fact, encourage participants to exert control.

Because it collapses across warning conditions, the interaction between trial type and WMC within the full quartile (2x2x4) ANOVA, wherein high spans in incongruent trials were significantly faster than lower-mid-spans, indicates that lower-mid-spans were the most rigid in their exertion of control. The similarity of high span RTs and low span RTs is most likely due to high spans' reliance on automaticity in the no-warning condition (as demonstrated by the error data) decreasing their average response latency. On the other end of the spectrum, the likeness of high span and upper-mid-span RTs may have resulted from the high end of the upper-mid-spans being somewhat flexible — or at least less rigid than true mid-spans — across warning conditions, decreasing the average RT of the quartile enough to attenuate previously-existing differences. Again, these data point to an increased ability of those with higher WMC to implement strategies for cognitive efficiency.

When comparing our response latency analyses of only the extreme upper and lower quartiles (2x2x2 ANOVA) to those from Kane and Engle (2003; mentioned above), it is important to note that their study included neutral trials (e.g., **TRUCK**), as well as differing congruency ratios within participants (0% congruent vs. 75% congruent). For the present purpose, we focus our discussion of their paper on the 75% congruent variation only. Kane and Engle's RT analyses investigated facilitation and inhibition, and their data yielded no difference across WMC for interference effects — that is, high and low spans produced similar response times during incongruent trials. In contrast, it appeared that facilitation effects (the latency difference between neutral and congruent trials) were greater for low spans, suggesting that low spans were more likely than high spans to use the information provided by the color words in congruent trials.

Although the present study did not include neutral trials, and so cannot speak to facilitation effects, we found that interference effects differed between low and high spans in our more established analyses (2x2x2), and differed between high spans and lower-mid-spans in the full quartile ANOVA. The likely reason for finding no difference after adding intermediate spans was an increase in standard error (see Figure 4). Still, the differences observed between high and the two lower quartiles spans make for compelling evidence that lower WMC leads to increased rigidity of control exertion. And these data are especially persuasive when evaluated alongside patterns shown by the Stroop error data.

As a whole, the change in error rate patterns across warning conditions implies that those with high WMC possess a unique ability to factor contextual cues into strategic conservation of cognitive resources. This increase is cognitive efficiency seems to be due to the amount of cognitive resources available to high spans (or high spans' ability to appropriately allocate resources) and sets them apart from the rest of the WMC spectrum. The dissimilarities in response latency analyses were most apparent between high spans and lower-mid-spans. These results add greatly to our understanding of the relationship between WMC and control exertion, because it displays a contrast between WMC levels within the range of individuals able to successfully exert control, compared to low spans. This allows for a more nuanced understanding of where and why changes in WMC produce varied response patterns in high-conflict paradigms.

Though the data are encouraging, we still have much to learn about the relationship between executive functioning and cognitive control. We acknowledge that the Stroop is anomalous in many ways. Important for us to keep in mind is that word reading — though still a learned task — is so well-practiced that we cannot consider it representative of learned automatic processes, as a whole. It might be worthwhile in future experiments to train participants on a novel task to the point of automaticity and then present task, which places a premium on a conflicting response. Given the that Stroop effects are best observed in paradigms employing a high-congruency ratio, determining whether differing levels of automaticity (e.g., automatic motoric, highly practiced learned automatic, and recently-learned automatic) alter control exertion across differing congruency ratios might allow for more comprehensive theories to be established.

Still, we believe our data create a strong case for understanding the control dilemma as a domain-free phenomenon. What's more, the data offer good indication that the inclusion of intermediate WMC individuals in analyses affords us greater theoretical leverage. In summary, we can confidently say that the error and response time data provide converging evidence, which suggests that in control dilemma scenarios, high spans are capable of maximizing cognitive efficiency. This efficiency manifests itself as an increase in the flexible exertion of cognitive control across varying contexts.

REFERENCES

Conway, A. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354-373. doi:10.1037/0096-3445.123.4.354

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross, B. H. Ross (Eds.), *The psychology of learning and motivation: Advances in research and theory, Vol. 44* (pp. 145-199). New York, NY, US: Elsevier Science.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York, NY, US: Cambridge University Press

Goschke, T. (2003). Voluntary action and cognitive control from a cognitive neuroscience perspective. In S. Maasen, W. Prinz, & G. Roth (Eds.), *Voluntary action: Brains, minds and sociality* (pp.49-85). New York, NY, US: Oxford University Press.

Hiebel, N., & Zimmer, H. D. (2015). Individual differences in working memory capacity and attentional control. *Canadian Journal of Experimental Psychology/Revu Canadienne De Psychologie Expérimentale*, *69*(1), 17-27. doi:10.1037/cep0000029

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY, US: Farrar, Straus and Giroux.

Kane, M. J., Conway, A. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. A. Conway, C. Jarrold, M. J. Kane, J. N. Towse, A. A. Conway, C. Jarrold, ... J. N. Towse (Eds.), *Variation in working memory* (pp. 21-46). New York, NY, US: Oxford University Press.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47-70. doi:10.1037/0096-3445.132.1.4

Kennette, L. N., & Wurm, L. H. (2014). On the disambiguation of meaning and the effect of cognitive load. *Current Psychology: A Journal For Diverse Perspectives On Diverse Psychological Issues*, doi:10.1007/s12144-014-9294-6

Lambert, A.E., Seegmiller, J.K., Stefanucci, J.K., & Watson, J.M. (2013). On working memory capacity and implicit associations between advanced age and dangerous driving stereotypes. *Applied Cognitive Psychology*, *27*(3), 306-313. doi:10.1002/acp.2908

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*,*35*(1), 196-204. doi:10.1037/a0014104

Meiring, L., Subramoney, S., Thomas, K. F., Decety, J., & Fourie, M. M. (2014). Empathy and helping: Effects of racial group membership and cognitive load. *South African Journal of Psychology*, *44*(4), 426-438. doi:10.1177/0081246314530280

Miller, A.E., Watson, J.M., & Strayer, D.L. (2012). Individual differences in working memory capacity predict action monitoring and the error-related negativity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*(3), 757-763. doi: 10.1037/a0026595

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY, US: Cambridge University Press.

Prakash, R. S., Hussain, M. A., & Schirda, B. (2015). The role of emotion regulation and cognitive control in the association between mindfulness disposition and stress. *Psychology and Aging*, *30*(1), 160-171. doi:10.1037/a0038544

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., and Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. PLoS ONE, *8*(1), e54402. doi:10.1371/journal.pone.0054402

Seegmiller, J.K., Watson, J.M., & Strayer, D.L. (2011). Individual differences in susceptibility to inattentional blindness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37,* 785-791. doi:10.1037/a0022474

Shipstead, Z., Lindsey, D.B., Marshall, R.L., & Engle, R.W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72116*-141. doi:10.1016/j.jml.2014.01.004

Stawarczyk, D., Majerus, S., Catale, C., & D'Argembeau, A. (2014). Relationships between wind-wandering and attentional control abilities in young adults and adolsecents. *Acta Psychologica*, *148*25-36. doi:10.1016/j.actpsy.2014.01.007

Storbeck, J., Davidson, N. A., Dahl, C. F., Blass, S., & Yung, E. (2015). Emotion, working memory task demands and individual differences predict behavior, cognitive effort and negative affect. *Cognition and Emotion*, *29*(1), 95-117. doi:10.1080/02699931.2014.9-04222

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498-505. doi:10.3758/BF03-192720

van Steenbergen, H. (2015). Affective modulation of cognitive control: A biobehavioral perspective. In G. E. Gendolla, M. Tops, S. L. Koole, G. E. Gendolla, M. Tops, & S. L. Koole (Eds.), *Handbook of biobehavioral approaches to self-regulation* (pp. 89-107). New York, NY, US: Springer Science and Business Media. doi:10.1007/978-1-4939-1236-0_7

Ward, A., & Mann, T. (2000). Don't mind if I do: Disinhibited eating under cognitive load. *Journal of Personality and Social Psychology*, *78*(4), 753-763. doi:10.1037/0022-3514.78.4.753

Watson, J.M., Bunting, M.F., Poole, B.J., & Conway, A. A. (2005) Individual differences in susceptibility to false memory in the Deese-Roediger-McDermott paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 76-85. doi:10.1037-/0278-7393.31.1.76

Watson, J. M., McDermott, K. B., & Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition*, *32*(1), 135-141. doi:10.3758/BF03195826

Watson, J.M., Miller, A.E., Lambert, A.E., & Strayer, D.L. (2008). *Simon says show me your tail: Individual differences in working memory capacity and cognitive control.* Paper presented at the 49th annual meeting of the Psychonomic Society, Chicago, Illinois.

Watson, J.M., Miller, A.E., Moffitt, C.C., Strayer, D.L, & Lambert, A.E. (2013). *To control or not to control? Using individual differences in working memory capacity to explore automation and the resolution of the control dilemma*. Paper presented at the 54th annual meeting of the Psychonomic Society, Toronto, Canada.